

# Northumbria Research Link

Citation: Younus, Othman, Hassan, Navid Bani, Ghassemlooy, Zabih, Haigh, Paul Anthony, Zvanovec, Stanislav, Alves, Luis Nero and Le Minh, Hoa (2020) Data Rate Enhancement in Optical Camera Communications using an Artificial Neural Network Equaliser. IEEE Access, 8. pp. 42656-42665. ISSN 2169-3536

Published by: IEEE

URL: <https://doi.org/10.1109/access.2020.2976537>  
<<https://doi.org/10.1109/access.2020.2976537>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/42369/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

# Data Rate Enhancement in Optical Camera Communications using an Artificial Neural Network Equaliser

OTHMAN ISAM YOUNUS<sup>1</sup>, NAVID BANI HASSAN<sup>1</sup>, ZABIH GHASSEMLOOY<sup>1</sup>, PAUL ANTHONY HAIGH<sup>2</sup>, STANISLAV ZVANOVEC<sup>3</sup>, LUIS NERO ALVES<sup>4</sup>, HOA LE MINH<sup>1</sup>

<sup>1</sup>Optical Communications Research Group, Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne NE1 8ST, U.K.

<sup>2</sup>Intelligent Sensing and Communications Group, School of Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.

<sup>3</sup>Department of Electromagnetic Field, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague 16627, Czech Republic

<sup>4</sup>Department of Electronics Telecommunications and Informatics, Instituto de Telecomunicações, University of Aveiro, Aveiro, Portugal

Corresponding author: Othman Isam Younus (e-mail: othman.younus@northumbria.ac.uk).

**ABSTRACT** In optical camera communication (OCC) systems leverage on the use of commercial off-the-shelf image sensors to perceive the spatial and temporal variation of light intensity to enable data transmission. However, the transmission data rate is mainly limited by the exposure time and the frame rate of the camera. In addition, the camera's sampling will introduce intersymbol interference (ISI), which will degrade the system performance. In this paper, an artificial neural network (ANN)-based equaliser with the adaptive algorithm is employed for the first time in the field of OCC to mitigate ISI and therefore increase the data rate. Unlike other communication systems, training of the ANN network in OCC is done only once in a lifetime for a range of different exposure time and the network can be stored with a look-up table. The proposed system is theoretically investigated and experimentally evaluated. The results record the highest bit rate for OCC using a single LED source and the Manchester line code (MLC) non-return to zero (NRZ) encoded signal. It also demonstrates 2 to 9 times improved bandwidth depending on the exposure times where the system's bit error rate is below the forward error correction limit.

**INDEX TERMS** optical camera communication, ANN equaliser, visible light communications, rolling shutter.

## I. INTRODUCTION

THERE are two type of optical receivers commonly used in visible light communication (VLC) systems: (i) photodiodes (PDs); and (ii) image sensors (ISs) [1-3]. The former is low cost and has higher bandwidth (or large detection areas). The latter, which is composed of a large number of PDs (i.e., pixels) aligned orthogonally, offer both multi-input multi-output capabilities and large detection areas but with complex data processing. Advances made in handheld smart devices have triggered the use of off-the-shelf conventional complementary metal oxide-semiconductor (CMOS) ISs for cameras. These devices can be used as an inherently integrated receiver (Rx) module in optical camera communications (OCC), which can capture light signals from a range of sources (i.e., LED-based traffic light, signage, headlights, vehicle tail lights, etc.) [3]. Moreover, in IS-based VLC systems the signal to noise ratio is independent of the transmission distance [4]. Note, as long as the projected image of the

transmitting LEDs cover a number of pixels, the incident light power level per pixel remains unchanged. However, the drawbacks of ISs-based Rx are low transmission bandwidth due to the camera frame rate  $R_f$  limitations and higher costs compared with PDs [5]. The maximum data rates of conventional commercial cameras are relatively low within a range of a few kbps, which are sufficient for non-data communications related applications such as device to device communications, sensing, and health care [6-13].

In [6], a data rate of 15 bit/s was demonstrated using an under-sampled frequency shift on-off keying (UFSOOK) and a 30-frames-per-second (fps) camera. The data rate increased to 150 bps using under-sampled phase shift on-off keying (UPSOOK), two LEDs, and a 50-fps camera [7]. In [8], a data rate of 150 bps was achieved using a red-green-blue (RGB) LEDs and a 50-fps in an IS-based VLC link. Alternatively, mobile-phone based VLC with a beacon jointed packet transmission scheme with rolling shutter (RS) was deployed with

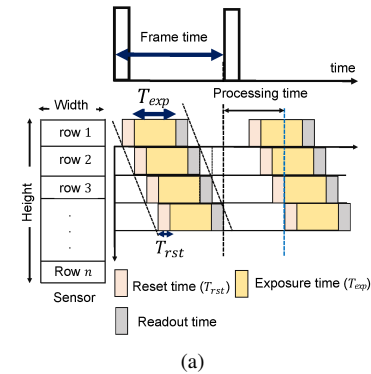
a data rate and  $R_f$  of 10.3 kbps and 60 fps for a link distance of 20 cm [9].

A colour-shift keying (CSK) modulation scheme was adopted for RGB LEDs based VLC and OCC offering data rates of 240 bps, 5.2 kbps and 8.64 kbps (i.e., 288 bit/frame with a  $R_f$  of 30 fps), respectively [10, 11, 13]. In [12], both data rate and packet size enhancement were employed in IS-based communications with 4-pulse amplitude modulation (PAM) to achieve a data rate of 3.6 kbps. Whereas in [14, 15], an OCC link with RS-based multilevel phase and amplitude modulation with the LED lights having spatial luminance distributions was investigated with data rates of 80, 78.5 bps for 256- and 64-QPSK, respectively with  $R_f$  of 30 fps. The use of equalisation techniques is adopted to compensate spatial and time dispersions in the OCC system. In [16], a double-equalisation was investigated for a single white LED. The 1<sup>st</sup> equaliser used a frame averaging based signal tracing (FAST) algorithm to extract the signal with nearly constant grayscale, while the 2<sup>nd</sup> scheme is an adaptive least mean square (LMS)-based feedforward equalizer (FFE), where the data rate achieved was 14.37 kb/s at  $R_f$  of 60 fps.

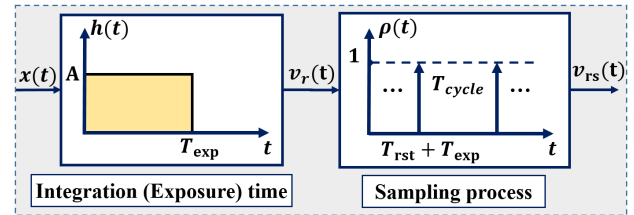
Additionally, different approaches can be adopted to improve the data rates in OCC with CMOS-based ISs including (i) high-speed cameras, which are highly costly and with limited applications; (ii) multiple transmitters, which may suffer from flickering [17]; and (iii) a special IS with a built-in PIN PD array has been used to increase the data rate to 55 Mbps using an optical orthogonal frequency division multiplexing [4]. However, the fabrication process of this IS is too complex and not commercially available. In contrast, the RS technique offers high-speed data transfer and has been widely adopted in OCC [18-21]. The concept of a 9.6 kbps VLC link using multiple transmitters (Tx) i.e., an array of  $8 \times 8$  LEDs and an IS of a resolution of  $320 \times 240$  with  $R_f$  of 30 fps was adopted in [22]. However, providing high capacities still imposes a number of challenges, which needs addressing [23, 24], including sampling rates, frame resolution image distortion factor and detection speed of the transmitted light source.

The sampling rate corresponds to exposure time  $T_{exp}$  i.e., the time where IS is exposed to the light. Increasing  $T_{exp}$  allows more lights to be integrated by the IS and increases the noise level, which is the key parameter in determining the bandwidth (i.e.,  $T_{exp}$  acts as a temporal low-pass filter (LPF) where it takes an average over a period of time. If the changes in the scene are faster than the bandwidth of the camera induced by the exposure time, the scene will be blurred).

ANNs as universal classifiers are used in VLC systems, the use of ANN is deployed for each channel in multiple-input multiple-output (MIMO) VLC [25]. In [26], an ANN-based equaliser was deployed in a fully connected mode to reduce the effect of the inter-symbol interference in a non-LOS (NLOS) or diffuse VLC link in indoor environment. It is also deployed in OCC systems to compensate for the data loss by reducing the gap-time between observed frames [28].



(a)



(b)

**FIGURE 1.** (a) An example of the frame acquisition based on rolling-shutter CMOS sensor. (b) The LSI model of a CMOS image sensor receiver.

Unlike previous studies, in this paper, for the first time, we propose and implement an ANN-based equaliser to mitigate the ISI induced by high sampling duration within the observed frame in OCC, and hence increase the transmission data rate.

The ANN network is trained only once for a range of  $T_{exp}$  and can be stored in a look-up table. We have developed an experimental test-bed for the proposed system for evaluating its performance in terms of the data rate, bit error rates (BER) and eye diagrams. In addition, we investigate a number of training methods and show that, the resilient back-propagation algorithm offers the best performance with a trained mean square error value of  $9.29 \times 10^{-5}$ .

We have also achieved the highest data rate in OCC using a single white LED source, the MLC NRZ encoded signal at the transmitter, an image sensor with  $R_f$  of 30 fps and an ANN-based equaliser at the receiver. The achievable bandwidth is also increased by approximately 9, 5, and 2 times for  $T_{exp}$  of 2, 1 and 0.5 ms, respectively compared with the existing reported systems [6].

The remainder of the paper is organised as follows: Section II introduces the OCC system's CMOS IS model. In Section III the proposed ANN-based post-equaliser is outlined, while in Section IV the system model is presented. The numerical results are presented and commented in Section V. Finally, section VI draws the final conclusion.

## II. CMOS IS MODELLING IN OCC SYSTEM

In CMOS ISs with RS, an array of pixels is used to capture the incident light in a progressive manner by exposing each row (column) of pixels as illustrated in Fig. 1(a). In RS-based cameras, the incident light at high frequencies and a relatively low exposure time is observed by forming different

illuminated bands indicating the “ON” or “OFF” status of the incident signal for OOK-NRZ at  $T_{\text{exp}}$ , or the intensity level at other modulation schemes, while light cannot preserve any signal during the resetting time  $T_{\text{rst}}$  and the readout time for a given row. The standard IS-based Rx is modelled as a linear shift-invariant (LSI) system, which is composed of two stages as depicted in Figure 1(b). The voltage at a pixel ( $U$ ,  $V$ ), which corresponds to an individual photodiode, is given by [27]

$$v_r(t) = \frac{A}{C_{\text{PD}}} \int_{t-T_{\text{exp}}}^t \Re. x(t) dt, \quad (1)$$

where  $A$  is the gain (set to unity for simplicity),  $C_{\text{PD}}$  and  $\Re$  are the equivalent capacitance and responsivity of the PD and  $x(t)$  is the received optical signal at the pixel ( $U$ ,  $V$ ) at time  $t$ . Note,  $T_{\text{exp}} \geq T_{\text{sym}}$ , where  $T_{\text{sym}}$  is the symbol period [29]. The number of symbols observed at the Rx depends on the resolution of the IS, exposure time, pixel clock, and the size of the region of interest [5]. The system response is given by [30]

$$h(t) = \frac{A}{C_{\text{PD}}} (u(t) - u(t - T_{\text{exp}})), \quad (2)$$

where  $u(t)$  is the unit step function. Integration of the input signal over  $T_{\text{exp}}$  results in a finite impulse response (FIR) LPF effect with a transfer function given as [30]

$$H(f) = \mathcal{F}\{h(t)\} = \frac{A T_{\text{exp}}}{C_{\text{PD}}} \frac{\sin(\pi f T_{\text{exp}})}{\pi f T_{\text{exp}}} e^{-j\pi f T_{\text{exp}}}. \quad (3)$$

The DC gain is proportional to  $T_{\text{exp}}$ , therefore a trade-off between the gain and the required bandwidth, where increasing  $T_{\text{exp}}$  will reduce the cut-off frequency. Following sampling, the discrete signal is given by [27]

$$v_{rs}(kT_{\text{cycle}}) = v_r(t) \sum_{n=-T}^T \delta(t - kT_{\text{cycle}}) + n(t), \quad (4)$$

where  $\delta(t)$  is the Dirac delta function,  $T_{\text{cycle}}$  is the sampling period and  $n(t)$  is the noise (i.e., signal-induced shot noise, ambient light induced shot noise and the thermal noise), which is modelled as zero-mean additive white Gaussian noise (AWGN).

### III. ARTIFICIAL NEURAL NETWORK EQUALISER

In RS-based OCC systems, the bandwidth limitation imposed by the sampling process of IS (i.e., LPF) results in ISI at higher data rates, thus leading to a significant degradation in system performance. The slow rise-time of the detected symbol is affected by the existence of the transition between different illumination levels, see Figure 2. Generally, a matched filter is adopted to mitigate ISI, however if this is not sufficient equalisation can be used to enhance data rates by estimation and mitigating the ISI effect [31]. Equalisation can be viewed from (i) the information theory, where ISI is predicted by the training filter coefficients based on a training sequence in order to minimise the error cost; and (ii) classification, where class decision boundaries are created in

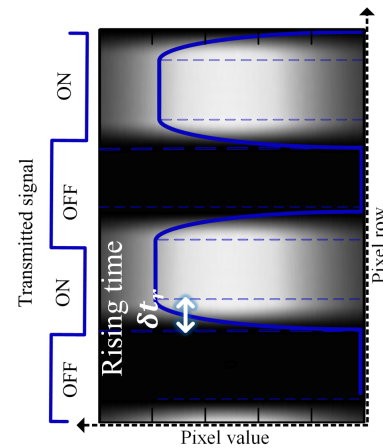
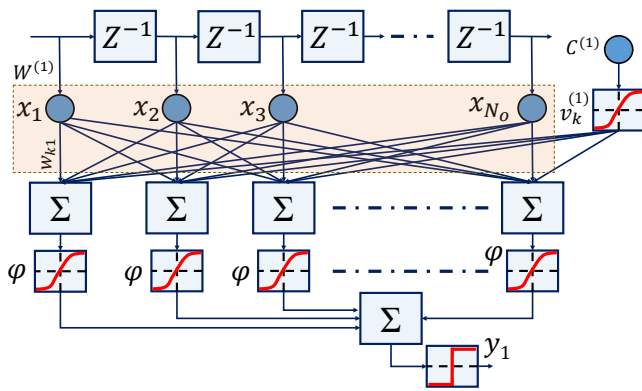


FIGURE 2. An example of the frame acquisition based on the RS CMOS sensor.

order to classify symbols based on training. The key difference between the two is that the former allows generalisation because of the use of boundaries, where unknown symbol transitions can be tolerated.

Linear decision boundaries are not sufficient to provide an optimal decision in practical channels, where the threshold boundaries are nonlinear. Therefore, ANN-based equalisers with the realisation of nonlinear decision boundaries offer improved performance in communication systems [25]. The classification and regression models are categorized as a subsection of supervised machine learning. The main difference between regression and the classification is that the output variable of the former is numerical (or continuous) whereas is categorical (or discrete) for the latter one [32]. Note, the boundaries have a high dependency on the number of neurons and the hidden layers, which are analogous to the human brain, where the synaptic weight is changed based on the training sequence. Different ANN approaches can be deployed for equalisation including the single-layer [33] and the multi-layer perceptron (MLP) [31]. In [26] it was shown that, MLP offers superior performance in mitigating ISI in optical wireless systems and hence, has been adopted in this work.

The ANN-based equaliser includes input layer, variable number of hidden layers ( $M-1$ ) and output layer. The neuron ( $k$ ) in the MLP filter in layer ( $m$ ) receives  $N_{m-1}$  tapped delay inputs  $[o_1^{(m-1)}, o_2^{(m-1)}, \dots, o_{N_{m-1}}^{(m-1)}]$ , known as the observation vector. The  $N_{m-1}$  tapped delay inputs are considered to study the impact of  $N$  previous samples on the desired sample. The  $n^{\text{th}}$  input has a corresponding weight connected to the  $k^{\text{th}}$  neuron  $w_{kn}^{(m)}$ . Note,  $N = T_{\text{exp}}/T_{\text{sample}}$ , where  $T_{\text{sample}}$  is the sampling duration which depends mainly on the pixel clock. The neurons represent the functional unit of the ANN, and are represented by a transformed version of the summed weighted inputs. The output is added to a constant offset value  $C^{(m)}$ , which is weighted with the threshold factor  $v_k^{(m)}$  to estimate the output  $o_k^{(m)}$  of the neuron  $k$



**FIGURE 3.** Block diagram of a single-hidden-layer ANN based equaliser. using a non-linear function  $f(\cdot)$ , given by [31]

$$o_k^{(m)} = f \left( \sum_{n=1}^{N_{m-1}} w_{kn}^{(m)} o_n^{(m-1)} + C^{(m)} v_k^{(m)} \right). \quad (5)$$

By considering a  $N_o \times 1$  input vector, and a  $N_M \times 1$  output vector the estimation of the next observation vector  $\mathbf{o}^{(m)}$  is required, where  $\mathbf{o}^{(o)} = \mathbf{x}$ ,  $\mathbf{o}^{(M)} = \mathbf{y}$ , and given  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_{N_o}]^T$ , and  $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_{N_M}]^T$ . Therefore, we have

$$\mathbf{y} = f \left( \mathbf{W}^{(M)} \mathbf{o}^{(M-1)} + C^{(M)} \mathbf{v}^{(M)} \right). \quad (6)$$

The MLP record its trained information in the weights  $w_{kn}^{(m)}$  and in the threshold factors  $\mathbf{v}_n^{(m)}$ , since  $C^{(m)}$  is given as a constant for all layers (set as  $C^{(m)} = 1, m = 1, 2, \dots, M$ ).

With a suitable number of neurons, a single hidden layer ANN is recognised as a universal approximator. Figure 3 illustrates the block diagram of a single layer ANN-based equaliser. The non-linear hyperbolic tangent sigmoid  $\varphi$  activation function is selected in this work since it provides a wider range for faster learning and grading, given by [34]

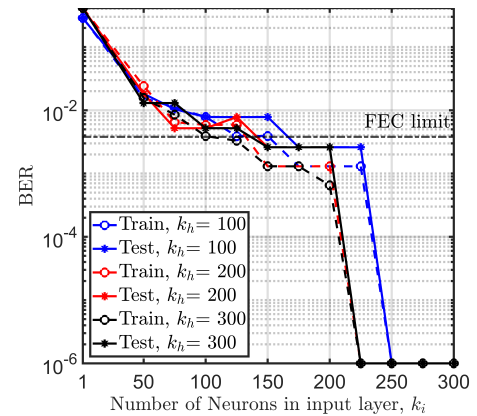
$$\varphi = \frac{e^{\beta\xi} - e^{-\beta\xi}}{e^{\beta\xi} + e^{-\beta\xi}}, \quad (7)$$

where  $\beta$  is the slope factor (the standard unity value is assigned) and  $\xi$  is the weighted input.

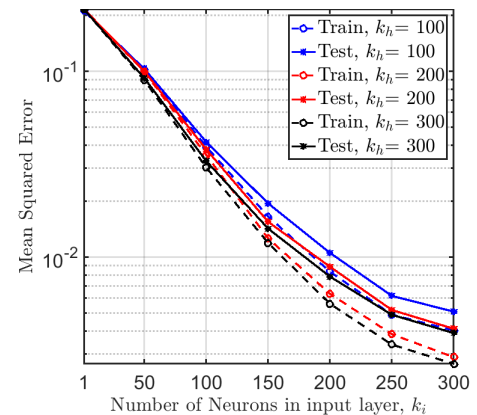
ANN can be trained using supervised and unsupervised methods [35, 36]. Resilient back-propagation (RBP) is an advanced version of back-propagation, which has been adopted to train ANN in this work. RBP, considered as one of the best learning methods in ANN [37], takes into account the sign of the error gradient to designate the direction of the weight update, and hence overcoming the slow convergence of the standard back-propagation algorithm and reducing the level of training compared with other algorithms. The concept of RBP is similar to the regular backpropagation technique, where ANN adjusts the weight in order to minimize the error cost function  $E_n$  as defined by [31]

$$E_k = ||d_k - y_k||^2 \quad (8)$$

where  $d_k$  is the ideal symbol value and  $y_k$  is the actual received value. RBP is an iterative operation, where the step



(a)



(b)

**FIGURE 4.** BER and MSE as a function of the number of neurons in input layer  $k_i$ , and the hidden layer  $k_h$  and  $N_{\text{bit}}$  of 192.

size is dynamically adapted for each weight depending on a gradient descent of  $E_k$ . The updated weight is given by [38]

$$w_{kn}(t+1) = w_{kn}(t) - \eta \frac{\partial E_k}{\partial w_{kn}}(t) + \mu \Delta w_{kn}(t-1), \quad (9)$$

where  $w_{kn}$  is the weight between the junction point of  $x_k$  and  $k^{\text{th}}$  neuron and  $\eta$  is the learning rate parameter.

The number of input neurons is set to be equal to the length of input patterns or vectors plus one, and the additional neuron being the bias neuron [39]. In the proposed system, the optimal number of neurons is selected based on the experimental measurements of the BER and the mean square error (MSE) values. Figure 4 shows the BER and MSE as a function of the number of neurons in the input and hidden layer. At the FEC BER limit, the number of neurons at the input and the hidden layers are 250 and 100, respectively which correspond to the MSE values of  $< 10^{-2}$  and  $6 \times 10^{-1}$ , respectively.

The momentum parameter  $\mu$  scales the impact of the previous step on the present one, thus introducing stability to the system and improving the convergence of the error function. Each weight has an individual evolving-value  $\Delta w_{kn}(t)$ , and the weight-step, which is only determined by its update-



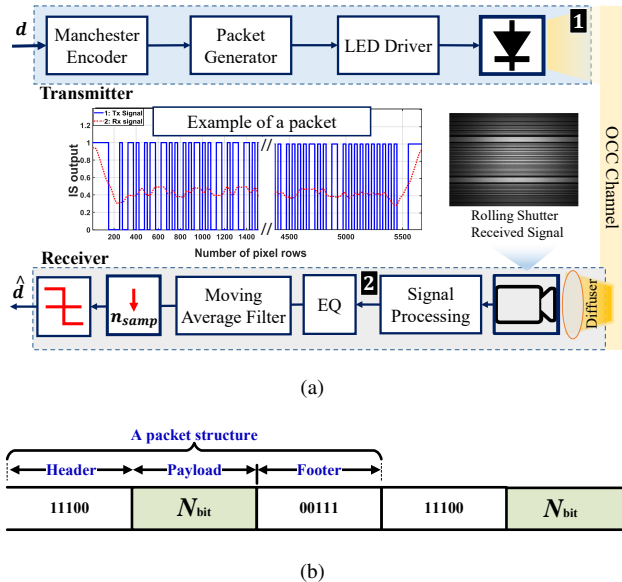


FIGURE 5. (a) System block diagram for OCC using LED and rolling shutter CMOS sensor. (b) The proposed structure of data frame packet.

value. The sign of the gradient is given by [38]

$$\Delta w_{kn}(t) = \begin{cases} +\Delta_{kn}(t), & \text{if } \frac{\partial E_k}{\partial w_{kn}}(k) > 0 \\ -\Delta_{kn}(t), & \text{if } \frac{\partial E_k}{\partial w_{kn}}(k) < 0 \\ 0, & \text{otherwise} \end{cases}, \quad (10)$$

where  $\Delta_{kn}(t)$  is the updated value. In order to provide a unity efficiency, the number of neurons is set to be similar to the number of tapped inputs [31].

#### IV. SYSTEM MODEL

The system block diagram of the proposed OCC system is illustrated in Figure 5(a). The transmitter is comprised of a pseudorandom binary sequence (PRBS)  $d$  with a length of  $2^{14} - 1$  is generated in MATLAB, which is then up-sampled and encoded using the non-return to zero (NRZ)-Manchester line code (MLC) in order to avoid flickering and a unity amplitude rectangular pulse shaping filter. The MLC signal, which ensures a uniform distribution of 1 and 0 symbols and facilitates both decoding and synchronisation processes of the signal [29], is packetized in order to ensure proper detection at the Rx as illustrated in Figure 5(b). Each packet consists of 5-bit header [11100],  $N_{\text{bit}}$ -bit payload and 5-bit footer [00111], which are used for intensity modulation of the LED via the optical driver.

The pattern of header and footer is designed such that (i) it never occurs in the MLC pattern for the payload; and (ii) the transition from one packet to another is smooth, which facilitates the training of the network. The transmitted optical signal over a free space channel is captured using a CMOS IS Rx (i.e., camera). The link distance is set to 60 cm, which can be extended by increasing the intensity of light, having a clear available region of interest and using the lenses at the transmitter and the receiver. For example, increasing the

optical transmit power from 1 to 4 mw will increase the transmission range by 100% [1].

A diffuser is utilised to distribute the captured light over the IS (i.e., the LED foot print was projected into the IS).

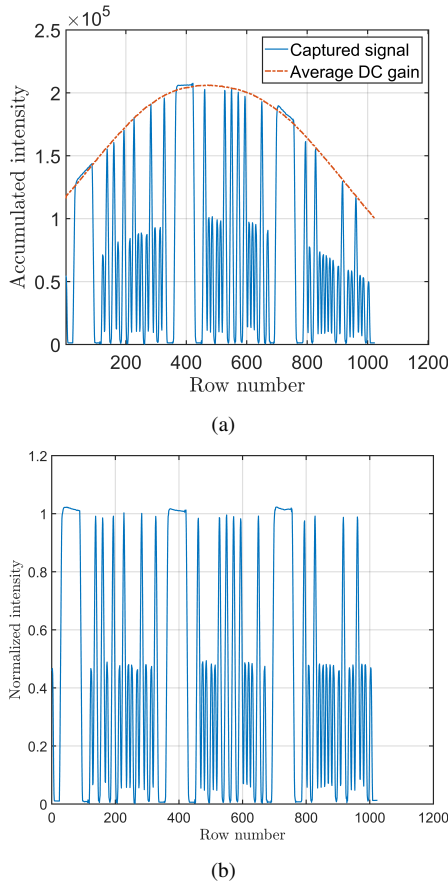
#### Algorithm 1: Signal extraction algorithm

**Input:** RS captured image  $P_{U \times V \times 3}$  and 20 frame of illumination gain (DC signal only)  $G_{U \times V \times 3}$

**Output:** The received signal is ordered the same as the transmitted signal

- 1 **foreach** Set of captured data frames  $P_{U \times V \times 3}$  **do**
- 2   Read  $U \times V \times 3$  sized colour plaintext image  $P_{U \times V \times 3} = [P(i, j, c)]$ . Let  $F = U \times V$ , and denote RGB components of  $P_{U \times V \times 3}$  as  $RP_{U \times V} = [RP(i, j)]$ ,  $GP_{U \times V} = [GP(i, j)]$  and  $BP_{U \times V} = [BP(i, j)]$ , respectively. Where  $i=1, 2, \dots, U, j=1, 2, \dots, V, c=1, 2, 3$ .
- 3   Convert the monochrome image to the grayscale image by calibrating RGB components  $RP_{U \times V}$ ,  $GP_{U \times V}$  and  $BP_{U \times V}$  together over 3, resulting  $PS_{U \times V}$ .
- 4   Accumulate intensities for all pixels at each row  $z = (z_i)_{i=1}^V$  where  $z_i = \sum_{j=1}^V PS_{U \times V}$ .
- 5   Calibrate  $z$  with respect to the averaged DC value  $z_{cal} = z / \overline{z_{DC}}$  where  $\overline{z_{DC}} = \sum_{j=1}^V GS_{i \times j}$  where  $\sum$  is the averaged DC.
- 6   Resample  $z_{cal}$  with respect to the packet length.
- 7   Locate the start of each packet in the frame.
- 8   Received packets are placed randomly within the captured frame due to the lack of synchronization between Tx and Rx, hence a correlation algorithm is adopted as follows: Create a filtered version of the ideal signal using a moving average filter  $z[s] = \frac{1}{M} \sum_{m=0}^{M-1} y[s+m]$  where  $M$  is the window size (set to  $n_{\text{samp}}$ ). Correlate the received signal with the filtered signal to find the correct sequence and change accordingly.

The RS received data is formed by accumulating intensities for all pixels at each row, since they are exposed to incident light at the same time. This increases the SNR of the signal by a maximum of  $V$  times, where  $V$  is the number of pixels in each row. The received signal strength forms the amplitude of the signal as explained in section II. The observed frames at the IS are first processed in MATLAB, where conversion to the grayscale is applied in order to eliminate the hue and saturation information while retaining the luminance of the image plane. Due to the non-uniformity of the illumination, the DC gain of the optical signal estimated from (3) is then measured to generate a



**FIGURE 6.** The received signal at a signal bandwidth of 400 Hz with  $T_{exp}$  of 1 ms: (a) without DC gain normalization and (b) with normalization.

calibration matrix. In order to mitigate the impact of noise, the calibration matrix is constructed by averaging over 20 frames of plain illumination, i.e., no AC signal. Note, none of the pixels of the region of interest in the calibration matrix should be over/under exposed, see Figures 6(a) and 6(b). The signal extraction algorithm is shown in Algorithm 1.

The frequency responses of the IS are first introduced for the conventional  $T_{exp}$  values of 0.5, 1 and 2 ms, which are also estimated using (3), where the frequency of the main lobe and the followed side-lobes is introduced in the following section. The main lobe indicates bandwidth available by means of 3 dB region (i.e., the cut-off frequency), whereas followed harmonics describe the rest of the magnitude spectrum. The harmonic frequency  $f_{sh}$  is used to specify the required bandwidth of the transmitted signal to ensure a minimum attenuation (i.e., at the peak of the side-lobes), where  $h$  is the harmonic number. Subsequently, a set of different signals are generated where  $N_{bit}$ -payload is varied with respect to  $f_{sh}$  in a way that ensures the capturing of three packets at every frame.  $T_{exp}$  is set to 2 ms and the values of selected  $N_{bit}$  are shown in Table 1. In a typical scenario, the bandwidth of the transmitted signal should be less than or equal to the cut-off frequency  $f_c$  as highlighted in Section II, whereas  $f_{sh}$  includes successive harmonics; The processed data is then applied to the proposed equaliser for different

Parameter	Value
Image sensor (IS)	DCC1645C
Maximum SNR of IS	44 dB
Exposure time $T_{exp}$	0.5, 1, and 2 ms
Camera raw image resolution	1280 × 1024
Pixel clock	10 MHz
Camera frame rate $R_f$	30 fps
Link distance	60 cm
LED type	SR-01-WC310
Number of sample $n_{samp}$	50
Number of bit per frame $N_{bit}$	108 to 1170 bits
Activation function	Hyperbolic tangent sigmoid
Number of neurons in input layer	128 ( $N_{bit} = 108$ bits), 250 ( $N_{bit} = 192$ bits), and 1200 ( $N_{bit} = 1170$ bits)
Number of neurons in output layer	1
Number of neurons in hidden layer	200
Number of hidden layer	1
Percentage of train to test	0.8
Maximum epochs	1000
Learning rate parameter $\eta$	0.01
Target BER	$10^{-4}$
Network training function	Resilient backpropagation Bayesian Regularization Polak–Ribière Conjugate Descent Scaled Conjugate Descent Conjugate Descent with Momentum

**TABLE 1.** System parameters.

$f_{sh}$ , and the quality of the transmitted signal is investigated in terms of the BER and the data rate.

A moving average filter (i.e., comparing the energy of current per chip with the previous one) is employed to reduce the noise while retaining a sharp response by means of averaging the input signal  $y$  in order to produce a discrete signal (i.e., point) given by [40]

$$z[s] = \frac{1}{L} \sum_{l=0}^{L-1} y[s + l], \quad (11)$$

where  $L$  is the number of points.

The data is then down-sampled prior to thresholding to recover the estimated version transmitted signal. Finally, we have adopted different training algorithms in order to achieve the optimal results in terms of the BER and the convergence speed. For this, we have used 80% of the data in the first frame for training, which corresponds to  $N_{bit}$ , and the number of bits used for training and testing is set based on  $N_{bit}$  per frame. For instance, for  $N_{bit}$  of 1170 bits, 80% (i.e., 936 bits) of the data is set to train the network while the rest (i.e., 20%, 234 bits) is used for consecutive testing purposes. Training is repeated every 10-  $N_{bit}$  since a set of 10 frames is considered for this purpose.

## V. RESULTS AND DISCUSSION

The system parameters used are shown in Table 1. Figure 7 shows the measured and simulated IS frequency response for a range of  $T_{exp}$ . The measured (and simulated) 3-dB  $f_c$  are 811 (1020), 443 (443) and 250 (250) Hz for  $T_{exp}$  of 0.5, 1 and 2 ms, respectively. We have used the odd frequencies  $f_{sh}$  (i.e., 716, 1231, 1736 and 2240 Hz, see Figure 7 for generating

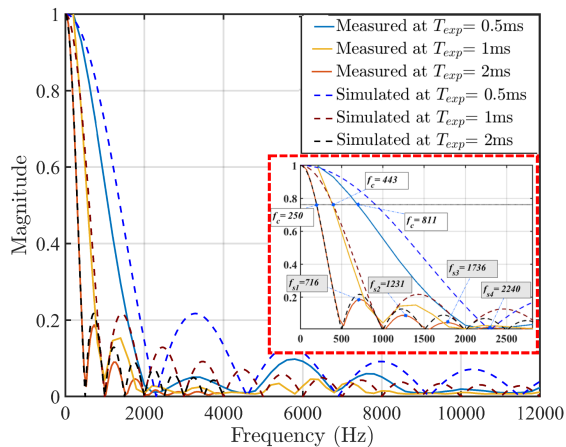


FIGURE 7. Image Sensor frequency response with different exposure time.

MLC data formats). Note, we have used an average of 636 bps per frame, thus a total of 12 kbps for 30 frames in order to meet the forward error correction (FEC) BER limit of  $3.8 \times 10^{-3}$ .

Next, the effective data rate of the system is evaluated following the methodology adopted in the previous section. The higher sampling rate is associated with the higher data rates used in this work. The BER as a function of the data rate for the system is initially measured prior to introduction of the proposed ANN equaliser as depicted in Figure 8. The proposed system demonstrates BER values below the FEC limit of  $3.8 \times 10^{-3}$  up to a data rate of 12 kbps for all three values of  $T_{exp}$  and a transmitted signal bandwidth of 2.275 kHz, which is around 9 times (for  $T_{exp}$  of 2 ms) higher the  $f_c$  of the unequalised system.

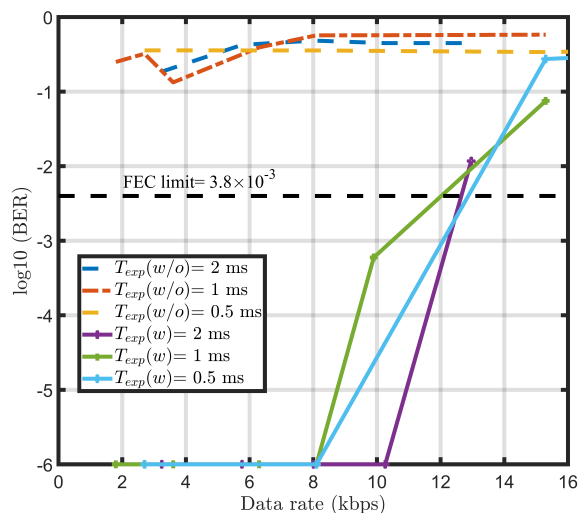


FIGURE 8. BER measurements of the system with and without equalisation in respect to effective data rates of different exposure times with  $R_f$  of 30.

The eye diagrams at  $T_{exp}$  of 2 ms without an ANN equaliser is illustrated in Figure 9(a), where the received signal shows three levels of amplitude in spite of sending an OOK signal due to the ISI, which creates an additional

level due to the transition delay in the status of the captured signal as explained in the previous section (see Figure 2). However, with an ANN equaliser the eye diagram depicted in Figure 9(b) shows significantly reduced ISI. Finally, we have evaluated the proposed system behaviour over time to determine the convergence time for different training algorithms for the hyperbolic tangent sigmoid activation function as illustrated in 10. The system performance is measured based on the estimated MSE values, which is estimated between the equaliser outputs and the desired outputs. The ideal case is considered, where the channel is assumed to be noise-free, and hence the error in the equaliser outputs is solely due to the channel dispersion. As shown, the resilient backpropagation algorithm displays faster convergence compared with others and offers the best performance (the estimated trained mean square error value is recorded as  $9.29 \times 10^{-5}$ ). The superiority of this algorithm is due to magnitudes elimination of the partial derivatives where the sign of the derivative is merely utilised to estimate the direction of the weight update.

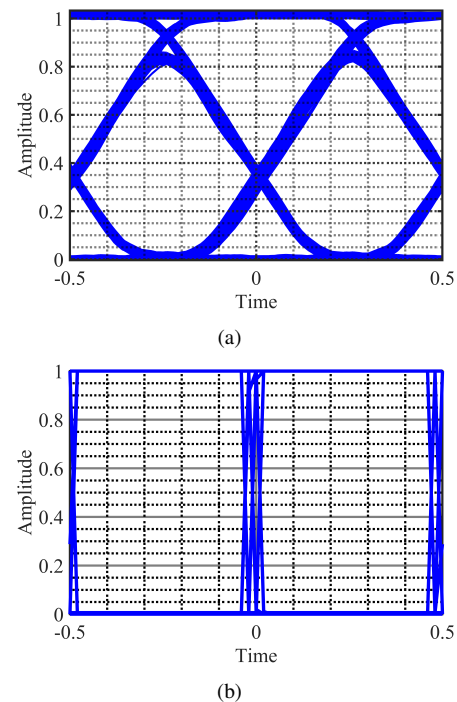
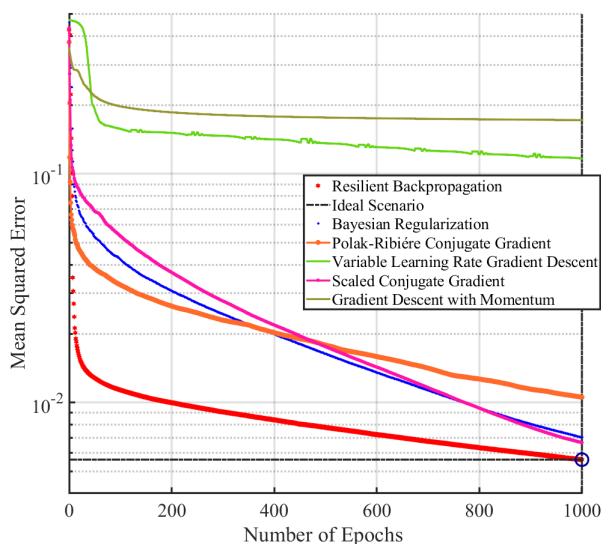


FIGURE 9. Eye diagram of the received signal at a signal bandwidth of 400 Hz of (a) without deploying the ANN-based receiver. (b) After applying the proposed receiver.

## VI. CONCLUSION

In this paper we for the first time demonstrate ANN adaptive equaliser at the receiver within the OCC system. The study provided the performance indicators for the proposed system. The data rates achieved were the highest in the OCC field, recorded as 12 kbps at the exposure time of 2, 1, and 0.5 ms using a single source and the MLC-NRZ encoded signal. The proposed system demonstrated the capability in retrieving the transmitted information with a bandwidth beyond the cut-





**FIGURE 10.** ANN system behaviour over time to calculate the faster convergence of different training algorithms for hyperbolic tangent sigmoid activation function.

off frequency limitation. Hence, it provided the bandwidth improvement of around 9, 5, and 2 times at the exposure times of 2, 1, and 0.5 ms, respectively.

## ACKNOWLEDGMENT

This work is supported by Northumbria University, the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no 764461 (VISION) and the UK EPSRC grant MARVEL (EP/P006280/1).

## REFERENCES

- [1] Z. Ghassemlooy, W. Popoola, and S. Rajbhandari, Optical wireless communications: system and channel modelling with Matlab®. CRC Press, 2012.
- [2] Luo, Pengfei and Zhang, Min and Ghassemlooy, Zabih and Zvanovec, Stanislav and Feng, Shulan and Zhang, Philipp, "Undersampled-Based Modulation Schemes for Optical Camera Communications," *IEEE Communications Magazine*, Vol. 56, No. 2, 204-212, (2018).
- [3] Do, T. H. and Yoo, M., "An in-Depth Survey of Visible Light Communication Based Positioning Systems," *Sensors (Basel)*, Vol. 16, No. 5, (2016).
- [4] Goto, Y. and Takai, I. and Yamazato, T. and Okada, H. and Fujii, T. and Kawahito, S. and Arai, S. and Yendo, T. and Kamakura, K., "A New Automotive VLC System Using Optical Communication Image Sensor," *IEEE Photonics Journal*, Vol. 8, No. 3, 1-17, (2016).
- [5] Do, Trong-Hop and Yoo, Myungsik, "Performance Analysis of Visible Light Communication Using CMOS Sensors," *Sensors*, Vol. 16, No. 3, 309, (2016).
- [6] R. D. Roberts, "Undersampled frequency shift ON-OFF keying (UF-SOOK) for camera communications (CamCom)," in *Wireless and Optical Communication Conference*, 2013, pp. 645-648.
- [7] P. Luo, Z. Ghassemlooy, H. Le Minh, X. Tang, and H.-M. Tsai, "Undersampled phase shift ON-OFF keying for camera communication," presented at Sixth International Conference on Wireless Communications and Signal Processing (WCSP), 2014.
- [8] Luo, P. and Zhang, M. and Ghassemlooy, Z. and Minh, H. Le and Tsai, H. and Tang, X. and Png, L. C. and Han, D., "Experimental Demonstration of RGB LED-Based Optical Camera Communications," *IEEE Photonics Journal*, Vol. 7, No. 5, 1-12, (2015).
- [9] Wang, W. and Chow, C. and Chen, C. and Hsieh, H. and Chen, Y., "Beacon Jointed Packet Reconstruction Scheme for Mobile-Phone Based Visible Light Communications Using Rolling Shutter," *IEEE Photonics Journal*, Vol. 9, No. 6, 1-6, (2017).
- [10] Chen, S. and Chow, C., "Color-Shift Keying and Code-Division Multiple-Access Transmission for RGB-LED Visible Light Communications Using Mobile Phone Camera," *IEEE Photonics Journal*, Vol. 6, No. 6, 1-6, (2014).
- [11] Chen, H. and Wen, S. and Wang, X. and Liang, M. and Li, M. and Li, Q. and Liu, Y., "Color-Shift Keying for Optical Camera Communication Using a Rolling Shutter Mode," *IEEE Photonics Journal*, Vol. 11, No. 2, 1-8, (2019).
- [12] Nguyen, D. T. and Chae, Y. and Park, Y., "Enhancement of Data Rate and Packet Size in Image Sensor Communications by Employing Constant Power 4-PAM," *IEEE Access*, Vol. 6, 8000-8010, (2018).
- [13] P. Hu, P. H. Pathak, X. Feng, H. Fu, and P. Mohapatra, "ColorBars: increasing data rate of LED-to-camera communication using color shift keying," presented at the Proceedings of the ACM Conference on Emerging Networking Experiments and Technologies, Heidelberg, Germany, 2015.
- [14] W. Chujo and M. Kinoshita, "Rolling-Shutter-Based QPSK by Spatial Luminance Distribution for Optical Camera Communication," in *2018 IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1-6.
- [15] Rachim, V. P. and Chung, W., "Multilevel Intensity-Modulation for Rolling Shutter-Based Optical Camera Communication," *IEEE Photonics Technology Letters*, Vol. 30, No. 10, 903-906, (2018).
- [16] L. Liu, R. Deng and L. Chen, "Spatial and Time Dispersions Compensation With Double-Equalization for Optical Camera Communications," *IEEE Photonics Technology Letters*, vol. 31, No. 21, 1753-1756, (2019).
- [17] Huang, Wei and Tian, Peng and Xu, Zhengyuan, "Design and implementation of a real-time CIM-MIMO optical camera communication system," *Optics Express*, Vol. 24, No. 21, 24567-24579, (2016).
- [18] C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a CMOS camera sensor for visible light communication," in *2012 IEEE Globecom Workshops*, 2012, pp. 1244-1248.
- [19] J. Ferrandiz-Lahuerta, D. Camps-Mur, and J. Paradells-Aspas, "A Reliable Asynchronous Protocol for VLC Communications Based on the Rolling Shutter Effect," in *2015 IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1-6.
- [20] T. Sonoda, H. Nagahara, K. Endo, Y. Sugiyama, and R. Taniguchi, "High-speed imaging using CMOS image sensor with quasi pixel-wise exposure," in *2016 IEEE International Conference on Computational Photography (ICCP)*, 2016, pp. 1-11.
- [21] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta, "Luxapose Indoor Positioning with Mobile Phones and Visible Light," presented at the Proceedings of the annual international conference on Mobile computing and networking - MobiCom '14, 2014.
- [22] T. Zinda, K. Ito, and W. Chujo, "Rolling-Shutter-Based Optical Camera Communication Using Distributed LED Array," in *International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, 2018, pp. 1-4.
- [23] Bani Hassan, Navid and Ghassemlooy, Zabih and Zvanovec, Stanislav and Biagi, Mauro and Vegni, Anna Maria and Zhang, Min and Luo, Pengfei, "Non-Line-of-Sight MIMO Space-Time Division Multiplexing Visible Light Optical Camera Communications," *Journal of Lightwave Technology*, Vol. 37, No. 10, 2409-2417, (2019).
- [24] Yang, Y. and Luo, J., "Composite Amplitude-Shift Keying for Effective LED-Camera VLC," *IEEE Transactions on Mobile Computing*, 1-1, (2019).
- [25] P. A. Haigh et al., "A MIMO-ANN system for increasing data rates in organic visible light communications systems," in *2013 IEEE International Conference on Communications (ICC)*, 2013, pp. 5322-5327.
- [26] Rajbhandari, Sujun and Faith, Joe and Ghassemlooy, Zabih and Angelova, Maia, "Comparative study of classifiers to mitigate intersymbol interference in diffuse indoor optical wireless communication links," *Optik*, Vol. 124, No. 20, 4192-4196, (2013).
- [27] X. Li, N. B. Hassan, A. Burton, Z. Ghassemlooy, S. Zvanovec, and R. Perez-Jimenez, "A Simplified Model for the Rolling Shutter Based Camera in Optical Camera Communications," presented at the *ConTEL*, Graz, Austria, 3-5 2019.
- [28] Liqiong Liu, Rui Deng, and Lian-Kuan Chen, "47-kbit/s RGB-LED-based optical camera communication based on 2D-CNN and XOR-based data loss compensation," *Opt. Express* 27, 33840-33846 (2019).

- [29] Nguyen, D. T. and Park, S. and Chae, Y. and Park, Y., "VLC/OCC Hybrid Optical Wireless Systems for Versatile Indoor Applications," *IEEE Access*, Vol. 7, 22371-22376, (2019).
- [30] J. C. Chau and T. D. C. Little, "Analysis of CMOS active pixel sensors as linear shift-invariant receivers," in *2015 IEEE International Conference*.
- [31] Haigh, P. A. and Ghassemloooy, Z. and Rajbhandari, S. and Papakonstantinou, I. and Popoola, W., "Visible Light Communications: 170 Mb/s Using an Artificial Neural Network Equalizer in a Low Bandwidth White Light Configuration," *Journal of Lightwave Technology*, Vol. 32, No. 9, 1807-1813, (2014).
- [32] H. Brink, J. Richards, and M. Fetherolf, *Real-world machine learning*. Manning Publications Co., 2016.
- [33] Lu, Xingyu and Lu, Chao and Yu, Weixiang and Qiao, Liang and Liang, Shangyu and Lau, Alan Pak Tao and Chi, Nan, "Memory-controlled deep LSTM neural network post-equalizer used in high-speed PAM VLC system," *Optics Express*, Vol. 27, No. 5, 7822-7833, (2019).
- [34] Zamanloo, B. and Mirhassani, M., "Efficient VLSI Implementation of Neural Networks With Hyperbolic Tangent Activation Function," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 22, No. 1, 39-48, (2014).
- [35] Rajbhandari, S. and Ghassemloooy, Z. and Angelova, M., "Effective Denoising and Adaptive Equalization of Indoor Optical Wireless Channel With Artificial Light Using the Discrete Wavelet Transform and Artificial Neural Network," *Journal of Lightwave Technology*, Vol. 27, No. 20, 4493-4500, (2009).
- [36] Teli, Shivani Rajendra and Zvanovec, Stanislav and Ghassemloooy, Zabihi, "Performance evaluation of neural network assisted motion detection schemes implemented within indoor optical camera based communications," *Optics Express*, Vol. 27, No. 17, 24082-24092, (2019).
- [37] Saputra, Widodo and Tulus and Zarlis, Muhammad and Sembiring, Rahmat Widia and Hartama, Dedy, "Analysis Resilient Algorithm on Artificial Neural Network Backpropagation," *Journal of Physics: Conference Series*, Vol. 930, 012035, (2017).
- [38] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," in *IEEE International Conference on Neural Networks*, 1993, pp. 586-591 vol.1.
- [39] Kak, Subhash. "Unary coding for neural network learning." *arXiv preprint arXiv:1009.4495*, (2010).
- [40] Loukas, A. and Simonetto, A. and Leus, G., "Distributed Autoregressive Moving Average Graph Filters," *IEEE Signal Processing Letters*, Vol. 22, No. 11, 1931-1935, (2015).



OTHMAN ISAM YOUNUS received the B.Sc. degree in Electrical Engineering/Electronics and Communication from the University of Mosul, Iraq, in 2013, and the M.Sc. degree in Microelectronics and Communication Engineering from Northumbria University, UK, in 2019. He is currently pursuing Ph.D. degree in Electrical Engineering at Northumbria University, UK. He has a 3 years' industrial experience as an RF and quality engineer, where he has participated in 4G

network deployment in Iraq as part of Alcatel-Lucent and Regional Telecom project. The project was awarded the best Middle Eastern project (Capacity Magazine, 2015).



NAVID BANI HASSAN received the B.Sc. degree in electrical engineering from Ferdowsi University of Mashhad, Iran, in 2012, and the M.Sc. degree from Tarbiat Modares University, Tehran, Iran, in 2015. He received the PhD degree from the University of Northumbria at Newcastle, UK, in 2019. He joined the Institute of Photonics at the University of Strathclyde, Glasgow, UK, in 2019 as Post-Doctoral Research Associate. His research interests include single-photon avalanche

detector arrays, micro-LED arrays, optical camera communications, and digital signal processing.



ZABIH GHASSEMLOOY CEng, Fellow of IET and OSA, Senior Member of IEEE: received BSc (Hons.) in electrical and electronics engineering from Manchester Metropolitan Univ., UK in 1981, and MSc (1984) and PhD (1987) from Univ. of Manchester, UK. 1987 to 1988 was a Post-Doctoral Research Fellow, City Univ., UK. 2004-2014 Northumbria, Newcastle as an Associate Dean for research in School of Engineering, and Faculty of Engineering and Environment. Currently is Head of Optical Communications Research Group. Research Fellow (2016) and Distinguished Professor (2015) at Chinese Academy of Science, Quanzhou. Published 825 papers (328 journals and 8 books), 97 keynote/invited talks and supervised 64-PhDs. Research interests: optical wireless communications, free space optics, visible light communications, radio over fibre/FSO. He is the Chief Editor of the British J. of Applied Science and Technology and the International J. of Optics and Applications, Associate Editor of a number of international journals, and Co-guest Editor of a number of special issues. Vice-Chair of OSA Technical Group of Optics in Digital Systems (2018-), Chair of the IEEE Student Branch at Northumbria University, Newcastle (2019-), was IEEE UK/IR Communications Chapter Secretary (2004-06), Vice-Chairman (2006-08), Chairman from (2008-11), and the Chairman of the IET Northumbria Network (2011-15).



PAUL ANTHONY HAIGH received his BEng and PhD degrees from Northumbria University, Newcastle upon Tyne, UK, in 2010 and 2014, respectively. Between 2011 and 2012, he undertook the prestigious Marie Curie Fellowship at the European Fellowship for Nuclear Research (CERN). During his PhD, he pioneered the topic of organic visible light communications with colleagues from UCL and Siemens AG, where he improved data rates in ultralow bandwidth organic photonic devices from kb/s to 55 Mb/s. Between 2014-2016 he was a Research Associate within the High Performance Networks Group at the University of Bristol, working on reconfigurable agile network interfaces. He is currently a Senior Research Associate at University College London working on high speed polymer visible light communications systems, with particular focus on modulation, digital signal processing and healthcare applications. He has published over 60 scholarly papers since 2012 including numerous articles in prestigious journals such as IEEE Journal on Selected Areas in Communications and IEEE Communications Magazine. He has collected >650 citations and a h-index of 16 to date.



STANISLAV ZVANOVEC (S'03-M'05) received the M.Sc. and Ph.D. degrees from Czech Technical University (CTU) in Prague in 2002 and 2006, respectively. He is currently a Full Professor and the Vice Head with the Department of Electromagnetic Field, Faculty of Electrical Engineering, CTU. He leads the Free-Space and Fiber Optics Team, CTU, and was the Vice Chair of WG1 of EU COST Project IC1101 OPTICWISE—Optical Wireless Communications—An Emerging Technology. His current research interests include wireless optical communications, visible light communications, and optical fiber sensors.



**LUIS NERO ALVES** (M'03) Graduated in 1996 and received his MSc degree in 2000, both in Electronics and Telecommunication Engineering from the University of Aveiro. Obtained the PhD degree in Electrical Engineering from the University of Aveiro in 2008. His expertise is on integrated circuit design for optical wireless applications. Since 2008 he has been the lead researcher from the Integrated Circuits group at Instituto de Telecomunicações, Aveiro. His current research

interest are associated to the design and performance analysis of optical wireless communication systems, with special flavor in visible light communications. Other research topics include materials and devices for IoT sensing devices. He has authored and coauthored several research articles in these fields. Luis Nero Alves has participated/coordinated several research projects in the fields of optical wireless communications and sensing devices (with both national funding – FCT/VIDAS, FCT/EECCO, IT/VLCLighting - and international funding – EU-CIP/LITES, EU-FP7/RTMGear, EU-COST/OPTICWISE, EU-COST/MEMOCIS, EU-H2020/MSCA/VisIoN, EU-H2020/FET/NeuralStimSpinal). He has also served in the technical program committee of several international conferences on the field of optical wireless communications. He has also served as reviewer for several journals in Optical communications, amongst which, IEEE/PTL, IEEE-OSA/JLT and Elsevier/OC.



**HOA LE MINH** is an Associate Professor (Reader) in Communications Engineering and the Director of Learning and Teaching for Maths, Physics and Electrical Engineering at Northumbria University, UK (2010 – present). Previously, he was a research fellow at University of Oxford, UK (2007-2010), research assistant at Siemens AG, Munich, Germany (2002-2004) and a lecturer at Ho Chi Minh University of Technology, Vietnam (1999-2001). He is an expert in photonics,

optical communications, visible light communications, Smartphone technology, signal processing and intelligent networks. He has published over 200 articles and book chapters in journals and conferences. He is the editor for a number of journals, and currently is the Chair of IEEE ComSoc UK and Ireland.

...